Difference of Stock Return Distributions and the Cross-Section of Expected Stock Returns

Preliminary, please do not quote.

Joon Chae* Wonse Kim** Eun Jung Lee***

This version: August 2, 2015

Abstract

We investigate the significance of difference of distributions (DD) over time in the cross-sectional pricing of stocks. Our noble proxy for DD, based upon the Earth Mover's Distance or the Wasserstein metric, measures the difference of empirical distributions of a firm's present stock return and those of its own past return. We find that stocks with higher DD exhibit higher returns on average, and the difference between returns on the portfolios with the highest and lowest DD is significantly positive. Moreover, the results from firm-level cross-sectional regressions show strong corroborating evidence for an economically and statistically significant positive relation between the DD and the expected stock returns. This positive relation persists after controlling for size, book-to-market, momentum, short-term reversals, liquidity, idiosyncratic volatility, skewness, kurtosis, and maximum return of the firm.

***Keywords*:** Difference of Distributions; Earth Mover's Distance; Wasserstein Metric; Expected Stock Returns

***JEL classfication*:** G11; G12; C14

* Graduate School of Business, Seoul National University, Seoul, Korea. Email: joonchae@snu.ac.kr
** Graduate School, Department of Mathematics, Seoul National University, Seoul, Korea. Email: acquinasws@gmail.com
*** College of Business Management and Economics, Hanyang University, Ansan, Korea. Email: ejunglee@hanyang.ac.kr
We are grateful to Kyoungwon Seo, Jae Hoon Hahn, Young Ho Eom, Noolee Kim, and other seminar participants at KAIST, UNIST, Hanyang University, Korea University, and Seoul National University for many helpful comments.

1. Introduction

In one of the most seminal papers in financial economics, Markowitz (1952) argues that there are two stages when we select a portfolio: firstly forming beliefs about asset returns and secondly optimizing our portfolio based upon the beliefs.[1] Traditional asset pricing models like the CAPM overlook the first stage and are constructed based upon the optimization in the second stage. Those models assume that investors already have beliefs about asset returns and know the form of distribution. For example, the CAPM assumes that asset returns follow a multivariate normal distribution or investors have quadratic utility function and that investors are mean-variance optimizing. However, empirical evidence confirms that portfolio returns are not normally distributed (Fama, 1965; Rosenberg, 1974), and even vague agreement about a specific stock return distribution does not exist (Tsay, 2010). That is, empirical evidence seems to suggest that investors do not know the distributional form of future stock returns.[2] In particular, a recent paper by Kacperczyk and Damien (2011) assumes that the form of the distribution of returns is not known, and proposes a novel method to incorporate "*distribution uncertainty*", uncertainty about the type of return distribution, to obtain an optimal portfolio. While the apparent difficulties of understanding the form of return distribution are generally recognized, surprisingly little is known about how seriously knowledge about

[1] Markowitz (1952) says that *"The process of selecting a portfolio may be divided into two stages. The first stage starts with observation and experience and ends with beliefs about the future performances of available securities. The second stage starts with the relevant beliefs about future performances and ends with the choice of portfolio. This paper is concerned with the second stage."*

[2] For example, Pastor and Stambaugh (2003), Ang and Bekaert (2004), and Guidolin and Timmermann (2007) show that distributions of assets returns tend to switch between different regimes. Liu et al. (2003) and Liu et al. (2005) suggest that the presence of rare events may perturb beliefs about the form of distributions. Welch (2000) argues that investors tend to differ in their assessment of future returns.

distribution affects empirical phenomena in finance such as the cross-sectional difference of asset returns. Therefore, in this paper, we empirically investigate whether there is a significant relation between the investors' understanding about return distributions and the expected stock returns.

To estimate investors' difficulty in understanding distributions of stock returns, we measure the stability of a stock's return distributions over the periods. If a distribution of stock returns in a current period is different than that from a previous period, we assume that investors will face more difficult problems in specifying a distribution of stock returns. To measure the difference of stock return distributions between the previous and present period, we adopt a noble metric widely used in the fields of mathematics and engineering. The earth mover's distance (EMD), the Wasserstein metric, describes the cost of moving one pile of dirt unto the other.[3] The cost of moving will be dependent on the distance between two piles as well as the difference in the shapes of the piles. In our case, each distribution can be regarded as a pile of dirt and the EMD measures the cost of transforming one distribution to the other. The EMD is not a unique measure to depict the difference among distributions, but it is considered to be a measure with more power.[4] We construct a measure of difference of distributions (DD) by averaging the EMDs estimated by using the previous 36-month daily stock return.

We examine the relation between the DD and cross-sectional pricing of stocks. To do this, we first sort stocks by the DD estimated in the previous 36 months and examine the subsequent monthly returns on the resulting portfolios over the period of 1965 to 2012.

[3] More thorough explanation about the EMD is in the following section.
[4] Please refer to the Appendix A for more details.

The results show that stocks with higher DD exhibit higher returns on average. The difference between value-weighted average returns on the portfolios with the highest and the lowest DD is around 0.72% in a month. The corresponding Fama-French-Carhart four-factor alpha from high-minus-low DD-sorted portfolios is 0.46% a month. We extensively investigate the robustness of our empirical results and find that the impact of the DD persists after accounting for firm characteristics, such as beta, size, book-to-market ratio, momentum, short-term reversal, and illiquidity. These results are also robust to control for return distribution characteristics, for instance, firms' idiosyncratic return volatility, skewness, kurtosis, and maximum daily return. Moreover, the effect of the DD on expected stock returns exhibit substantial persistence in firm-level cross-sectional regressions, even after controlling for a variety of other firm-level variables. Our empirical results suggest that investors are averse to high DD stocks and ask for a discount to buy those stocks, and thus, those stocks exhibit higher returns in the future.

Our empirical findings seem to be consistent with theories of ambiguity. Knight (1921) says, with ambiguity, the location and shape of the distribution is open to question. As in Ellsberg (1961) and the survey of Camerer and Weber (1992), ambiguity is generally defined as uncertainty about distribution. If an investor has to optimize her portfolio without knowing the distribution of stock returns, she is much like an agent in Ellsberg's paradox,[5] one of the most popular examples of ambiguity aversion. The preference for existence of a specific distribution as in Ellsberg's paradox is referred to as ambiguity aversion (Ellsberg, 1961; Sherman, 1974; and etc.). An agent with ambiguity aversion is

[5] In Ellsberg's paradox (Ellsberg, 1961), an agent mostly prefers a game with known probability distribution (for example, a game with an urn containing 50 red balls and 50 black balls) to one without a specific probability distribution (for instance, a game with an urn containing 100 balls of unknown number of red and black balls). For more thorough discussion about Ellsberg type example, see Ellsberg (1961), Epstein and Schneider (2008), and others.

more likely to dislike her situation that she can barely specify a stock return distribution since the distributions were too unstable in the past. If we interpret the size of the DD as the degree of ambiguity, we can argue that more ambiguous stocks will significantly outperform less ambiguous stock in the future. Actually, the DD can be directly linked to the definition of ambiguity, "uncertainty about distribution", by Knight (1921) or the explanation of ambiguity by Epstein and Schneider (2008). In their paper, they argue that when quality is difficult to judge, investors treat signals as ambiguous. Since investors need to spend more resources to understand a more unstable distribution of a stock[6], investors will feel more uncertain and ask for a premium to hold the stock. This result implies that there is a positive relation between ambiguity and the expected stock return as most theories about ambiguity suggest. Our empirical results show the evidence of a positive premium for bearing ambiguity. This positive premium is also helpful to understand the DD as a proxy for ambiguity. Even though various theoretical, experimental, and survey studies about ambiguity[7] argue that positive ambiguity premium exists, almost none of the empirical studies verify such a positive premium (Jiang et al., 2005; Zhang, 2006; Ehsani et al., 2013). For example, Baltussen, Bekkum, and Grient (2013) find a negative relation at a firm level by developing a proxy for ambiguity, volatility of implied volatility.[8] However, this negative relation is inconsistent with the general perception of ambiguity aversion and its premium as in Epstein and Schneider

[6] This way to understand ambiguity is closely related to the multi-prior approach by Garlappi, Uppal, and Wang (2007) even though their paper is more concentrated on parametric and model uncertainty. On the other hand, our paper deals with non-parametric characteristics of distribution uncertainty.

[7] See, among many others, Gilboa and Schmeidler (1989), Dow and Werlang (1992), Epstein and Wang (1994), Olsen and Troughton (2000), Klibanoff, Marinacci, and Mukerji (2005), Epstein and Schneider (2008), Ju and Miao (2012), Izhakian (2012), Drechsler (2013), Jeong, Kim and Park (2015), and Ahn, Choi, Gale and Kariv (2014)

[8] In an aggregate level, Brenner and Izhakian (2011) and Erbaş and Mirakhor (2007) find negative relation between ambiguity and market return, but Anderson et al. (2009) find positive relations by constructing a proxy for ambiguity, the dispersion of forecasts from the Survey of Professional Forecasters (SPF).

(2008), Klibanoff, Marinacci, and Mukerji (2005), and others. To summarize, many theories argue that ambiguity aversion makes investors ask for ambiguity premium in stock market, no empirical studies cross-sectionally confirm ambiguity premium in a stock level. On the other hand, the DD has statistically significant positive relation with future stock returns as most theories about ambiguity suggest.

An alternative interpretation of our empirical finding is consistent with literature of real options. Bulan (2005) argues that a firm with more real options, under an uncertain environment, is likely to delay the exercise of options, such as follow-up investment since delaying exercise of the options is more valuable for the firm. If we interpret the DD as a measure for uncertainty, a company with higher DD will have a higher value of real options under this uncertain environment, and less incentives for investing. In this case, the company will delay investment and its current stock price will be relatively discounted. As Grullon et al. (2012) shows in their paper, the effect of real options on stock return changes when the options are exercised. The exercise of the options, i.e., initiating investment, will make stock price reflect the value of a new investment, thus future stock return will increase. Our empirical findings that higher DD stocks exhibit higher future stock returns are consistent with real option theory.

The rest of this paper is organized as follows. In Section 2, we describe our dataset and construct variables of interest including the DD. Section 3 and Section 4 report our empirical results, and Section 5 provides discussions and interpretations. We conclude in Section 6.

2. Data and Construction of Variables

2.1 Data

The sample data include returns from the Center for Research on Security Prices (CRSP) Daily Stock File and book value from the Compustat of all stocks listed in NYSE, Amex, and NASDAQ. CRSP is used to obtain prices, daily returns, market returns, shares outstanding, trading volume, etc. We also obtain balance sheet information including assets, liabilities, and the total equity from Compustat. We use stock prices and shares outstanding to calculate market capitalization, and use daily returns to calculate DD for each firm in each month as well as beta, idiosyncratic volatility, skewness, kurtosis, and maximum daily return. These variables are defined in detail in the Appendix B. The sample period spans from January 1965 to December 2012. To be included in the final sample for a given month, at least 100 daily returns must exist in the previous 12 months.

2.2 Difference of Distributions (DD)

When an investor optimizes her investment and consumption, first she needs to specify the distribution of stock returns. Since most theoretical and empirical works in financial economics overlook the importance of this specification process, in this paper we investigate the role of this process. To verify the relation between investors' difficulty to specify a stock return distribution and the expected stock return, we define a measure that reflects investors' difficulty. We argue that investors have trouble deciding a form of stock return distribution if distributions change over time more considerably. Therefore, we introduce a statistical measure, DD aforementioned, representing variability of stock return distributions over time.

To construct a variable representing the difference of distributions, we apply the earth mover's distance (EMD) in computer science, also called the Wasserstein metric in mathematics. Informally, the EMD is defined as the minimum amount of work needed to change one pile of dirt into the other, where the amount of work is measured by quantity of dirt moved times the distance by which it is moved. With stock return data, the stock return distribution in a period can be considered to be a pile of dirt aforementioned. If a stock's return distribution in period t-2 is different than that in period t-1, the EMD of the stock will be larger, and we interpret this stock as one with which investors are hardly able to specify its return distribution. The EMD is not a unique measure to estimate the difference among distributions. There are other measures such as the Kolmogorov-Smirnov statistic or Kullback-Liebler divergence. However, the EMD has comparative advantages over other measures as explained in Appendix A. The formal definition of the EMD and its actual calculating procedures are also explained in detail in Appendix A.

In this paper, our measure of the difference of distributions (DD) is the average of 24 EMDs in the 2 years before the portfolio formation month. To estimate each EMD at t, we compare a stock's return distribution in t-12 to t-1 and that in t-11 to t. For example, the EMD in January 1999 is calculated by comparing the daily return distribution between January 1998 and December 1998, and distribution between February 1998 and January 1999. The same procedure is repeated until the 24th EMD in December 2000 is estimated. Finally, to construct the DD in January 2001, these 24 EMDs from January 1999 to December 2000 are averaged. By averaging 24 EMDs to construct a DD, we can effectively control outliers in return distributions and construct a measure that reflects distribution changes month by month.

2.3 Summary Statistics

Table 1 reports the summary statistics and correlation coefficients for variables of interest. The beta of a stock for a month (BETA) is estimated by regressing the daily stock return on the value weighted index return using a previous year sample. SIZE is the natural logarithm of the market value of equity of the company (in thousands of dollars) measured by times series average of a firm's market capitalization for the most recent 12 months. Book-to-market ratio (BM) is the book value of equity divided by its market value at the end of the last fiscal year. Momentum (MOM) is calculated as the return over the 11 months prior to that month. Short-term reversal (REV) is the previous month’s stock return. We use the illiquidity (ILLIQ) measure by Amihud (2002), which is measured as the ratio of the absolute monthly stock return to its dollar trading volume for each stock. Idiosyncratic volatility (IDIOVOL) is calculated as the standard deviation of the daily residuals in month t from the CAPM. We calculate skewness (SKEW) as the historical third-order centralized moment using daily returns within year t. Kurtosis (KURT) is defined by the fourth-order centralized moment. We define MAX as the maximum daily return over the past month.

Panel A in Table 1 presents the summary statistics of DD, beta, firm size, book-to-market ratio, momentum, short-term reversal, illiquidity, idiosyncratic volatility, skewness, kurtosis, and maximum return. The mean of DD is 0.0251. The lowest-percentile (P1) and the highest-percentile (P99) in the DD are 0.0159 and 0.0617, respectively. The mean and the median of SIZE are 11.37 and 11.25, respectively. The mean and median of BM are 0.8364 and 0.6581, indicating a right skewness in the distribution. The mean, the median, and the highest-percentile (P99) of MAX are 0.0742, 0.0504, and 0.4068, respectively. Panel B shows correlation coefficients of variables of interest. DD is negatively correlated with size, book-to market ratio, momentum, and short-term reversal, but positively correlated with beta and idiosyncratic volatility.

- Insert Table I about here –

2.4 Descriptive statistics based on Difference of Distributions

To get a clearer picture of the characteristics of the DD, we compute the statistics for decile portfolios of stocks based on DD. Table II presents the average values for DD, SIZE, BM BETA, MOM, REV, ILLIQ, IDIOVOL, SKEW, KURT, MAX, AGE, and REAL OPTION_D. Definitions of all variables are given in Appendix B.

The results show that there is a cross-sectional variation of the DDs; the DD increases from 0.0199 to 0.0407. First, the market capitalization and firm's age differences are substantial. The market capitalization and firm's age decrease quite dramatically for the highest DD deciles. It is not surprising since a newly-public firm value will be evaluated using the value of future investment opportunities set rather than assets in place, thus there is a higher chance to have a higher difference of stock return distribution. The stocks with the highest DD tend to have higher beta, suggesting that stocks with a high DD are more exposed to market risk. As we move from the lowest DD to the highest DD decile, book-to-market ratio increases. The column (8) in Table II reports the monotonic increase in idiosyncratic volatility as DD increases. It suggests that stocks with a high DD in a given month will have high realized idiosyncratic volatility in the same month measured using the residuals from a daily market model within the month. Meanwhile, the stocks with the highest DD tend to have higher maximum daily returns, indicating the maximum daily returns increases an average of 0.1440 for decile 10.

Since real options are more valuable for a firm with an uncertain environment as in Bulan (2005), a company with a higher DD, presumably under a more uncertain environment, will accumulate more real options. Therefore, we argue that a company with more real options must have a higher DD on average if the explanation with real options can be valid. Grullon et al. (2012) argue that high-tech, pharmaceutical, and biotechnology industries have plenty of real options. Thus, we use a dummy variable for real options (REAL OPTION_D) with the value of 1 if a firm belongs to high-tech, pharmaceutical, and biotechnology industries, or 0 otherwise. The final column of Table II shows that the stocks with the highest DD tend to accumulate more real options (REAL OPTION_D). We also test whether there is a cross-sectional variation of the DDs across industries. In an unreported table[9], we calculate average values of the DD for 10 different industries defined based on the 10 Fama-French industries. The results show considerable cross-sectional variation of the values of DD across industries and higher averages of DDs of high-tech (0.0280), or health-care, medical equipment, and drug (0.0279) industries than those of other industries (from 0.0244 to 0.0267). It confirms that companies with a higher DD seem to have more real options.

- Insert Table II about here –

3. Difference of Distributions and the Cross-Section of Expected Stock Returns

[9] For brevity, this information will be available upon request.

3.1 Univariate Analysis

The first empirical investigation is whether the DD can explain the cross-sectional variation of expected stock returns. Table III reports the value-weighted and equal-weighted average monthly returns of decile portfolios formed on the DD. To construct this table, we first calculate these measures for each sample firm over the previous month. Each month we sort stocks into 10 value-weighted/equal weighted portfolios using our measure for the DD. The portfolios sorted on the DD demonstrate strong variation in mean return, as shown in Table III. Taking a closer look at the value-weighted average returns across deciles, stocks (S) with the lowest DD provide 0.7960% of expected return per month on average and the stocks (B) with the highest DD do 1.5166%. It suggests that going from decile 1 to decile 10, value-weighted average returns increase significantly. The results show that the equal-weighted average returns on the decile portfolios sorted by the DD increase monotonically in portfolio rank. The bottom decile portfolio (S) by the DD has 1.0988% of the expected return per month on average and the top decile has 4.6755%. Overall, we find significant evidence that stocks with higher DDs have higher expected returns than stocks with lower DDs. This implies that since investors need to spend more resources to understand unfamiliar distributions of a stock compared to that of the stock in past years, they may require a premium to hold the stock. Our results show the evidence of a positive premium for bearing DD. It suggests that investors will face more difficult problems in specifying a distribution of stock returns as the DD increases.

- Insert Table III about here -

Table IV shows the value-weighted average returns of the different portfolios using various estimation and holding periods. The portfolios are formed based on *J*-month and held for *K*-months. We confirm that all the trading strategies with a long position in stocks with the highest DD and a short position in stocks with the lowest DD yield positive returns, and the individual t-statistics are sufficiently large to be significant. For example, the mean portfolio return for the 24-month/6-month strategy increases from 5.3913% in decile 1 (S) to 9.4979% in decile 10 (B). A trading strategy with a long position in B stocks and a short position in S stocks (B-S) yields a return of 4.1066% (t=6.71) per month. The more successful strategy selects stocks based on their returns over the previous 36 months. The mean portfolio return for the 36-month/12-month strategy increases from 10.61% in decile 1 to 21.03% in decile 10. This strategy yields 10.43% per month with *t*-statistic of 10.16. The evidence of higher returns for stocks with a high DD than for stocks with a low DD is robust to different estimations and holding periods.

- Insert Table IV about here -

3.2 Bivariate Analysis Sorted on Difference of Distributions and Firm Characteristics

We examine the relation between DD and future stock returns after controlling for firm characteristics. For example, stocks with high DD tend to be small and illiquid. In particular, following Olsen and Troughton (2000), 84% of respondents agreed that estimates of future stock return distributions are more unreliable for smaller firms than for larger firms. To ensure that the effect of the DD is not driven by these characteristics, we investigate the profitability of portfolios sorted by DD after controlling for firm characteristics, such as beta, size, book-to-market ratio, momentum, short-term reversal, and illiquidity.

Table V shows monthly returns averaged across the portfolios formed by dependent two-way sorts on a stock return's DD and firm characteristics, following Bali et al. (2011) and Baltussen et al. (2013).[10] First, stocks are categorized into 10 groups by firm characteristics. Then, within each decile portfolio, we further sort stocks into decile portfolios ranked based upon the DD, which results in a total of 100 portfolios. Next, we average each of the DD portfolios across the firm characteristic deciles. As Baltussen et al. (2013) argue, we can control for each firm characteristic without assuming a parametric form about the relation between the DD and the future stock returns. For each of these portfolios, we calculate the average equal-weighted/value-weighted returns over the following month.

The first column of Panel A in Table V reports the value-weighted returns averaged across the ten beta deciles to produce decile portfolios with dispersion in DD. Since we average across beta deciles, the produced decile portfolios sorted by DD will include all betas. The portfolio returns for each month are calculated as a value-weighted average of returns from strategies initiated at the end of the past month. After controlling for beta, the average return difference between the low and high DD portfolios is about 0.655% per month with a *t*-statistic of 3.63. It suggests that the positive relation between the DD and the future stock returns is not affected by beta. Column 2 in Panel A shows that the highest DD firms earn an average of 2.3133%, compared to 1.2468% for the smallest DD firms, when we control for size. The return differential between these two deciles (B-S)

[10] We find qualitatively similar results when we conduct independent bivariate sorts on the DD and firm characteristics.

is 1.0665% and significant (t=4.85). The results from two-way sorts on a stock return's DD and SIZE are larger and statistically more significant than those reported for the univariate sort in Table III. As shown in Table I, the high DD stocks are negatively correlated with the firm size, which suggests that these stocks should have higher returns. When controlling for book-to market ratio (BM), the return differentials between B and S are also positive and significant. When stocks are sorted based on momentum, the average return of the Big-Small portfolio is 0.6695%, with a *t*-statistic of 3.82. Subsequently, the average excess return of the B-S portfolio equals 0.72% per month when controlling for short-term reversal. Finally, we see whether the illiquidity explains the higher returns for the highest DD stocks relative to the lowest DD stocks. The average return of the B-S portfolio is 1.0046% per month with a *t*-statistic of 4.76. These results suggest that a positive DD premium remains economically important and firm characteristics do not explain the positive relation between DD and futures stock returns.

Panel B of Table V presents equal-weighted average monthly returns to portfolios formed by two-way sorts on the DD and firm characteristics. Similarly we find confirmatory evidence in Panel B with equal-weighted average monthly returns. After controlling for firm characteristics, the equal-weighted average return differences between the lowest DD and the highest DD portfolios are in the range of 1.37% - 3.11% per month with high significance.

These results suggest that for the value-weighted and the equal-weighted portfolios, the well-known cross-sectional effects such as beta, size, book-to-market ratio, momentum, short-term reversal, and illiquidity cannot explain the high returns of the high DD stocks.

- Insert Table V about here -

3.3 Alphas

In this section, we examine whether a rational risk-based approach can explain our result that the degree of DD provides premium. Table VI presents the value-weighted/equal-weighted portfolios' postranking alphas estimated under three different factor specifications, the capital asset pricing model (CAPM), the three-factor model proposed in Fama and French (1993), and the four-factor model proposed in Carhart (1997). The results in Panel A show that our measures for DD are highly correlated with alphas estimated from three different factor specifications. The magnitude of the alpha is positively related to the level of the DD, which implies that the high DD portfolios earn more positive abnormal returns. All three alphas of the B-S spread are significantly positive. The CAPM alpha is 0.4889% per month (t=2.27), the three-factor alpha is 0.3878% per month (t=2.40), and the four-factor alpha is 0.458% per month (t=2.85). A simple trading strategy of B-S portfolio generates about 6% of annual abnormal return, after controlling for the market, size, value, and momentum effects. This pattern of alphas from the three different factor specifications implies that the abnormal returns of B-S portfolios are not specific to asset pricing models and confirms our hypothesis of the DD premium.

As shown in the Panel B, similar and more economically and statistically significant results are obtained for the monthly returns on equal-weighted portfolios. The alphas difference between the lowest DD and the highest DD portfolios are in the range of 3.15% - 3.41% per month with significance. For example, the CAPM alpha of the B-S spread is 3.3064% per month (t=10.80) and the four-factor alpha is 3.4136% per month (t=12.86).

- Insert Table VI about here -

3.4 Firm-level Cross-sectional Regressions

We have examined the significance of the DD as a determinant of the expected stock returns at the portfolio level. However, as mentioned in Bali et al. (2011), the portfolio-level analysis has potentially significant disadvantages.[11] Therefore, we examine the cross-sectional relation between the DD and the expected stocks returns at the firm level using Fama-MacBeth (1973) regressions. The dependent variable is one-month ahead monthly return.

Table VII reports the time-series averages of the coefficients from the regressions of expected stock returns on the DD, beta, size, book-to-market ratio, momentum, short-term reversal, and illiquidity over the sample period, 1965 - 2012. These variables are defined in detail in Appendix B. In the univariate regression of expected return on the DD, the coefficient is positive and statistically significant. The average coefficient from the monthly regressions of expected return on the DD is 0.3851 with a *t*-statistic of 3.09. When the six control variables are included in the regression, the coefficient on the lagged DD remains large and significant. The results show that the coefficients on the six individual control variables are as expected. Of these six control variables, BM and REV contribute the most to the expected stock returns. The average coefficient on BM is significantly positive and that on REV is significantly negative, as expected. Overall, the results from cross-sectional regressions show strong corroborating evidence for an

[11] Bali et al. (2011) argue two potential disadvantages: "*First, it throws away a large amount of information in the cross-section via aggregation. Second, it is a difficult setting in which to control for multiple effects or factors simultaneously.*"

economically and statistically significant positive relation between the degree of the DD and the expected stock returns.

- Insert Table VII about here -

4. Difference of Distributions and Return Distribution Characteristics

4.1 Bivariate Analysis Sorted on Difference of Distributions and Return Distribution Characteristics

We examine the relation between the DD and the future stock returns after controlling for various return distribution characteristics, such as idiosyncratic volatility, skewness, kurtosis, and maximum return. Firstly, Epstein and Schneider (2008) propose that expected excess returns will increase with idiosyncratic volatility in fundamentals. On the other hand, Ang et al. (2006, 2009) empirically find that stocks with high idiosyncratic volatility have low subsequent returns. To verify that the DD is not explained by one of the return distribution characteristics, in this case idiosyncratic volatility, we examine the profitability of the DD after controlling for it. Secondly, we examine the profitability of the DD after controlling for skewness and kurtosis. Since extant empirical evidence indicates that skewness and kurtosis are related to the expected returns, we examine whether the profitability of the strategy is confined to the subsample of stocks based on these return distribution characteristics. For example, Arditti (1967), Levy (1969), Arditti and Levy (1975), and Kraus and Litzenberger (1976) extend the standard portfolio theory to incorporate the effect of skewness on valuation. Harvey and Siddique (2000) present an asset-pricing model with conditional co-skewness, where risk-averse investors prefer positively skewed assets to negatively skewed assets. Assets that decrease a portfolio's skewness (i.e., that make the portfolio returns more left-skewed) are less desirable and should command higher expected returns. Meanwhile, Fang and Lai (1997) note that investors are compensated for bearing kurtosis risk via excess returns. Dittmar (2002) extends the three-moment asset-pricing model using the restriction of decreasing absolute prudence (Pratt and Zeckhauser, 1987; Kimball, 1993). He argues that investors with decreasing absolute prudence dislike co-kurtosis, suggesting preference for lower kurtosis. Investors are averse to kurtosis, and prefer stocks with lower probability mass in the tails of the distribution to stocks with higher probability mass in the tails of the distribution. Assets that increase a portfolio's kurtosis (i.e, that make the portfolio returns more leptokurtic) are less desirable and should command higher expected returns. Therefore, we examine the relation between the DD and the future stock returns after controlling for skewness and kurtosis.

Thirdly, Bali et al. (2011) argue that there is a negative and significant relation between the maximum daily return over the past month and the future stock returns. Since this factor has power in explaining stock returns, and is related to the shape of the distribution of stock returns, we examine whether the maximum return can explain the positive DD effect.

To this end, we apply the same dependent two-way sorting procedure as in the previous section. We form decile portfolios at the end of each month by sorting the return distribution characteristics, such as the idiosyncratic volatility, skewness, kurtosis, and maximum returns. Then we sort each decile portfolio into ten additional DD portfolios. We further average each of the DD portfolios across the ten decile portfolios. Next, we form a B-S portfolio that buys the resulting highest DD portfolio and sells the resulting lowest DD portfolio. The portfolio returns for each month are calculated as an equal-weighted/value-weighted average of returns over the following month from strategies set up at the end of each previous month.

The results of these double sorts are presented in Table VIII. Panel A in Table VIII presents the value-weighted monthly returns across the return distribution characteristics. After controlling for idiosyncratic volatility (IDIOVOL), the value-weighted average monthly return difference between the lowest DD and the highest DD portfolios is about 0.4341% with a *t*-statistic of 4.66. The results show that the positive relation between DD and expected stock returns is not affected by idiosyncratic volatility. Since idiosyncratic volatility and the DD are highly correlated, the range of returns of DD portfolios is reduced after controlling for idiosyncratic volatility, but still significant. It suggests that idiosyncratic volatility does not explain the DD effect. We control for skewness and kurtosis in a similar way. We calculate the value-weighted average returns of portfolios formed by dependent two-way sorts on the DD and higher moments, such as skewness and kurtosis. When controlling for the magnitude of the skewness in the portfolio formation month, the highest DD firms earn an average of 1.6049% monthly, compared to 0.9038% for the lowest DD firms. The return differential between these two deciles (B-S) is 0.7012% and significant (*t*=3.32). When we control for kurtosis in column (3), the return difference between the lowest DD and the highest DD portfolios is 0.6103% per month and statistically significant at all conventional levels. These findings indicate that none of these higher moments are able to explain the DD effect. We control for the maximum daily return over the past month (MAX), with the results reported in the fourth column of

Table VIII, Panel A. The effect of the DD is preserved, with an equal-weighted average return difference (0.7210%) between the lowest DD and the highest DD portfolios and a corresponding *t*-statistic of 5.40.

Next, we turn to an examination of the equal-weighted average monthly returns on the DD portfolios after controlling for the return distribution characteristics as in Table VIII, Panel B. For the equal-weighted average return, the results are even stronger than those presented for the value-weighted average return in Panel A. The results show that the equal-weighted average return differences between the lowest DD and the highest DD portfolios are in the range of 1.1071% – 3.6453% per month, and they are statistically significant. Overall, a strategy of buying the highest DD firms and shorting the lowest DD firms seems to produce high returns when controlling for return distribution characteristics, such as idiosyncratic volatility, skewness, kurtosis, and maximum return.

- Insert Table VIII about here –

4.2 Regression Tests

Another way to examine the DD effect after controlling for return distribution characteristics is to look at firm-level cross-sectional regressions of expected stock returns on DD values. We conduct cross-sectional regressions of future stock returns on the DDs after controlling for various return distribution characteristics. Table IX presents the time-series averages of the coefficients from the regressions of future stock returns on the DD, idiosyncratic volatility, skewness, kurtosis, and maximum return as well as beta, size, book-to-market ratio, momentum, short-term reversal, and illiquidity over the sample period of 1965 - 2012. These variables are defined in detail in Appendix B. The results show that the DD effects reported earlier are still evident in this cross-sectional test. As in prior studies, we find a strong positive DD effect on returns, even after controlling for idiosyncratic volatility, skewness, kurtosis, and maximum return.

Harvey and Siddique (1999, 2000) and Smith (2007) distinguish systematic and idiosyncratic skewness from total skewness, and argue that stocks with lower systematic skewness outperform stocks with higher systematic skewness. Following Harvey and Siddique (2000), systematic skewness is the coefficient of a regression of returns on squared market returns, including the market return as a second regressor. The idiosyncratic skewness is the skewness of the residuals from this regression. The results show that when systematic skewness and idiosyncratic skewness are included in the regression, the coefficient on the DD remains large and significant. Overall, the positive DD effects persistently exist in firm-level cross-sectional regressions, even after controlling for a variety of other firm-level variables.

- Insert Table IX about here –

5. Discussions

A plausible interpretation of our evidence is that the DD may proxy for ambiguity. There are several possible reasons why we interpret the degree of DD as ambiguity. First, it is directly linked to the definition of ambiguity, "uncertainty about distribution", as explained by Knight (1921) and by Epstein and Schneider (2008). In Epstein and Schneider's paper, they assume that a signal s is the sum of a parameter of interest (θ) and noise of the signal (ε). They argue that when the variance of ε has a wider range, investors treat signals as ambiguous. For example, if the distribution of a signal (s_1) has a normal distribution of N(0, 1), N (0, 2), or N (0, 3) and that of a signal (s_2) has a normal distribution of N (0, 1), N (0, 4), or N (0, 9), they argue that the signal (s_2) is more ambiguous than the signal (s_1). According to this theoretical situation, as the range of variance of ε expands, the DD mentioned in our study should also increase. Based on these arguments, we can interpret the DD as a potential proxy for ambiguity.

Second, our results show the evidence of a positive premium for holding the stocks with a higher DD. Since Knight (1921) distinguishes uncertainty from risk, various studies have argued that Knightian uncertainty or ambiguity is important in explaining a firm's profit (Knight, 1921), economic decision making (Knight, 1921; Keynes, 1921), procyclical price-consumption and price-dividend ratio, and other economic phenomena (Ju and Miao, 2012). Most theoretical, experimental, and survey studies argue that positive ambiguity premiums exist (Jeong, Kim and Park, 2015; Ahn, Choi, Gale and Kariv, 2014; Olsen and Troughton, 2000)[12]. For example, Yates and Zukowski (1976) show that decision makers are willing to pay an average ambiguity premium of 20% of the expected value. On the other hand, only a few empirical studies exist[13] and most of them find negative relation between ambiguity and the expected returns. For instance, Zhang (2006) explains short-term price continuation by ambiguity. Baltussen, Bekkum, and Grient (2013) find negative relation between ambiguity and expected stock returns by developing a proxy for ambiguity, which is volatility of implied volatility. However, this negative relation is inconsistent with general perception of ambiguity aversion and its

[12] See, among many others, Gilboa and Schmeidler (1989), Dow and Werlang (1992), Ju and Miao (2012), Epstein and Schneider (2008), Klibanoff, Marinacci, and Mukerji (2005), Izhakian (2012), Epstein and Wang (1994), and Drechsler (2013).
[13] Ehsani et al., (2013), Brenner and Izhakian (2011), Anderson et al. (2009),

premium as in Epstein and Schneider (2008), Klibanoff, Marinacci and Mukerji (2005), and others. To summarize, many theories argue that ambiguity aversion makes investors require ambiguity premium in the stock market. No empirical studies cross-sectionally confirm ambiguity premium at a stock level. On the other hand, our simple proxy, free from any assumption about distribution, has a statistically significant positive relation with future stock returns, as most theories about ambiguity suggest. Even though we believe that our proxy is not the only one, nor necessarily the best one, to represent ambiguity, we argue that this paper suggests the interpretation of the DD as ambiguity.

One of the advantages of using the DD as the proxy of ambiguity is that it is easy to apply the proxy to research about the cross-section of every single stock return, since only daily stock returns are necessary to construct this proxy. Extant empirical studies such as Zhang (2006) and Jiang, Lee and Zhang (2005) use firm characteristic proxies, for example, firm size, age, analyst coverage, dispersion in analyst forecasts, return volatility, and cash flow volatility. In addition, Ehsani et al. (2013) and Baltussen et al. (2013) use option-implied volatility for their ambiguity proxy using data that have time series and corss-sectional limitation. On the other hand, our study only uses individual stock prices. Thus, by the virtue of the simplicity of our proxy, numerous implications from theories about ambiguity can be comprehensively investigated.

Another possible explanation of the results in this paper is related to literature on real options. Bulan (2005) argues that a firm with more real options under uncertain environment is likely to delay the exercise of options because delaying exercise of the options is more valuable for the firm. She finds empirical evidence that the more uncertain a firm's future cash flow is, the more valuable real options to delay investment there are. If we interpret the DD as a measure for uncertainty, a company with a higher DD will have higher values of real options under this uncertain environment, and less incentives for investing. In this case, the company will delay investment and its current stock price will be relatively discounted. As Grullon et al. (2012) show in their paper, the effect of real options on stock return changes when the options are exercised. The exercise of the options, i.e., initiating investment, will make the stock price reflect the value of the new investment, thus the future stock return will increase. Our empirical findings that higher DD stocks exhibit higher future stock returns are consistent with this real option theory.

Meanwhile, our results are close in spirit to the literature on parameter uncertainty and distribution uncertainty. Garlappi et al. (2007) develop a model to incorporate parametric and model uncertainty by using a multi-prior approach. To study parametric uncertainty, their framework incorporates the case that an investor estimates parameters, such as expected returns, using sample observations of the realized returns. However, Kacperczyk and Damien (2011) say that it is very difficult to reach any agreement about the precise form of the underlying stochastic process driving returns. Kacperczyk and Damien (2011) propose a concept of "distribution uncertainty", which discuss uncertainty about the type of return distribution. Our study, which is in the same vein as Kacperczyk and Damien (2011), non-parametrically measures distribution uncertainty by computing the difference of stock return distributions between the previous and the present period.

6. Conclusion

In this paper, we examine the relation between the DD and cross-sectional pricing of stocks. The results show that stocks with a higher DD exhibit higher returns on average. The difference between value-weighted average returns on the portfolios with the highest and the lowest DD is around 0.72% in a month. The corresponding Fama-French-Carhart four-factor alpha from high-minus-low DD-sorted portfolios is 0.46% a month. We extensively investigate the robustness of our empirical results and find that the impact of the DD persists after accounting for firm characteristics, such as beta, size, book-to-market ratio, momentum, short-term reversal, and illiquidity. These results are also robust to control for return distribution characteristics, for instance, firms' idiosyncratic return volatility, skewness, kurtosis, and maximum daily return. Moreover, the effect of the DD on expected stock returns exhibits substantial persistence in firm-level cross-sectional regressions, even after controlling for a variety of other firm-level variables. Our empirical results suggest that investors will face more difficult problems in specifying a distribution of stock returns as the DD increases. This paper suggests the interpretation of the degree of DD as the level of ambiguity. The notion of DD is consistent with the definition of ambiguity by Knight (1921). The positive relation between our proxy for ambiguity (the DD) and future stock returns is what most theories about ambiguity suggest. Almost none of the other empirical studies verify such a relation.

Appendix A. Introduction to Earth Mover's Distance (Wasserstein Metric)

In mathematics, various metrics quantify the distance between two probability distributions. Henceforth, we will use 'probability metric' to mean 'metric' measuring the distance between two probability distributions. The probability metric we adopt in this paper is the Earth Mover's Distance (EMD), the Wasserstein metric in mathematics. Intuitively, each distribution can be regarded as a unit amount of boxes piled on R. Then the EMD takes a minimum cost[14] of transforming one pile into the other.

Let X, Y be random variables with their induced probability measures P_X, P_Y. Then, the p-th EMD between two probability measures is defined by

$$W_P(P_X, P_X) \coloneqq \left(inf_{\gamma \in \Gamma(P_X, P_Y)} \int_{R \times R} d(x, y)^p d\gamma(x, y) \right)^{\frac{1}{p}}, \qquad \text{(A.1)}$$

Where $\Gamma(P_X, P_Y)$ denotes the set of all probability measures on $R \times R$ with marginal probability P_X and P_Y on the first and second factors respectively and $d(x, y)$ denote a metric defined In R, so we take absolute value of difference for the metric $d(x, y)$. Therefore, in this paper, the p-th EMD metric between two probability measures P_x, P_Y have special form,

$$W_P(P_X, P_X) \coloneqq \left(inf_{\gamma \in \Gamma(P_X, P_Y)} \int_{R \times R} |x - y|^p d\gamma(x, y) \right)^{\frac{1}{p}} \quad \text{(A.2)}$$

It is worth noting that the p in the formula (A.2) is a parameter we need to decide. Therefore, choosing the value p depends merely on each problem we face and we have to find out the optimal value of p.

---

14 In general, cost is defined as the amount of box needed to be moved times the distance to be moved

The EMD has two nice advantages. First, it is properly finer for the theories of statistics. That is, the EMD metricizes weak convergence.[15] Second, it is easy to compute. To compute the EMD in practice, we only focus on discrete probability measures. It means that normalized histogram[16] used as a discrete probability measure is a proxy for original probability measure.

Consider the discrete measures $P_X \coloneqq \sum_i P_i \delta_{x_i}$ and $P_Y \coloneqq \sum_j P_j \delta_{y_j}$. Then, to calculate the EMD, it is enough to solve a following problem of linear program.

$$\underset{(in\,\pi)}{\text{Minimize}} \sum_{i,j} \pi_{i,j} |x_i - y_i|^p (= W_P(P_X, P_Y)^p)$$

$$\text{Subject to} \sum_j \pi_{i,j} = p_i$$

$$\sum_i \pi_{i,j} = p_j$$

$$\pi_{i,j} \geq 0$$

15 (Definition.1) A sequence of probability measure P_n on R converges weakly to P(in symbol $P_n \Rightarrow P$), if

$$|\int h\, dP_n - \int h\, dP| \to 0, \text{ as } n \to 0$$

for all bounded, continuous functions h. This is exactly same notion with "convergence in distribution" which is one of the most importance convergence concepts in statistics.

(Definition.2) If a metric d has the property that

$$d(P_n, P) \to 0 \textit{ if and only if } P_n \Rightarrow P,$$

we say that d "metricizes weak convergence".

(Theorem) 1-th Wassersstein metric "metricizes weak convergence".

16 Since we use histogram to calculate the EMD, the problem of choosing the number and width of bins in a histogram does happen. In this paper, we use $\sqrt{n}$ as the number of bins in a histogram (n: the number of data.).

Appendix B. Variable Definitions

BETA: The beta of a stock for a month (BETA) is estimated by regressing the weekly stock return on the value weighted index return using a previous year sample.

SIZE: Firm size (SIZE) is measured by the time series average of market capitalization of a company for the most recent 12 months. In our regressions, we take the natural logarithm of size.

BM: Book-to-market ratio (BM) is book value of equity in the most recent fiscal year divided by market value of equity.

MOM: Following Jegadeesh and Titman (1993), momentum (MOM) is the cumulative stock return over the previous 11 months starting two months ago to isolate momentum from the short-term reversal effect.

REV: Following Jegadeesh (1990) and Lehmann (1990), we measure short-term reversal (REV) for each stock in month t as the return on the stock over the previous month.

ILLIQ: Following Amihud (2002), stock illiquidity (ILLIQ) is defined as the ratio of the absolute monthly stock return to its dollar trading volume.

IDIOVOL: We compute the monthly idiosyncratic volatility (IDIOVOL) for each stock *i* as the standard deviation of the daily residuals in month *t* from the CAPM. Specifically, we estimate $R_{i,d} - r_{f,d} = \alpha_i + \beta_i(R_{m,d} - r_{f,d}) + \varepsilon_{i,d}$, where $\varepsilon_{i,d}$ is the idiosyncratic return on day *d*. The idiosyncratic volatility of stock *i* in month *t* is defined as the standard deviation of daily residuals in month t: $\text{IVOL}_{i,t} = \sqrt{VAR(\varepsilon_{i,d})}$.

SKEW: The total skewness of stock i for month t is the historical third-order centralized moment using daily returns within year t: $\text{SKEW}_{i,t} = \frac{1}{D_t}\sum_{d=1}^{D_t}\left(\frac{R_{i,d}-\mu_i}{\sigma_i}\right)^3$, where D_t is the number of trading days in year t, $R_{i,t}$ is the return on stock i on day d, μ_i is the mean of returns of stock *i* in year *t*, and σ_i is the standard deviation of returns of stock *i* in year *t*.

SSKEW and ISKEW: Following Bali et al. (2011), we decompose total skewness into systematic and idiosyncratic components by estimating the following regression for each stock: $R_{i,d} - r_{f,d} = \alpha_i + \beta_i(R_{m,d} - r_{f,d}) + \gamma_i(R_{m,d} - r_{f,d})^2 + \varepsilon_{i,d}$, where $R_{i,d}$ is the return on stock i on day d, $R_{m,d}$ is the market return on day d, $r_{f,d}$ is the risk-free return on day d. The systematic skewness (SSKEW) of stock *i* in year *t* is the estimated slope coefficient $\hat{\gamma}_{i,t}$. The idiosyncratic skewness(ISKEW) of stock *i* in year *t* is defined as the skewness of daily residuals $\varepsilon_{i,d}$ in year *t*.

KURT: Kurtosis is the fourth-order centralized moment.

MAX: Following Bali et al. (2011), maximum return of each stock is the maximum daily return over the past month.

AGE: Firm age (AGE) is the number of years since the firm was first covered by CRSP.

REAL OPTION_D: REAL OPTION_D is a dummy variable for real options with the value of 1 if a firm belongs to high-tech, pharmaceutical, and biotechnology industries, or 0 otherwise.

References

Ahn, David, Syngjoo Choi, Douglas Gale, and Shachar Kariv, 2014, Estimating ambiguity aversion in a portfolio choice experiment, *Quantitative Economics* 5, 195-223.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31-56.

Anderson, Evan W., Eric Ghysels, and Jennifer L. Juergens, 2009, The impact of risk and uncertainty on expected returns, *Journal of Financial Economics* 94, 233–263.

Ang, Andrew and Geert Bekaert, 2004, How regimes affect asset allocation, *Financial Analysts Journal* 60, 86-99.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259-299.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2009, High idiosyncratic volatility and low returns: International and further US evidence, *Journal of Financial Economics* 91, 1–23.

Arditti, Fred D. and Haim Levy, 1975, Portfolio efficiency analysis in three moments: The multi period case, *Journal of Finance* 30, 797-809.

Arditti, Fred D., 1967, Risk and the required return on equity, *Journal of Finance* 22, 19-36.

Bali, Turan G., NusretCakici, and Robert F. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427-446.

Baltussen, Guido, Sjoerd Van Bekkum, and Bart Van Der Grient, 2013, Unknown unknowns: Vol-of-vol and the cross-section of stock returns, *Working Paper Erasmus University*.

Bernasconi, Michele, and Graham Loomes, 1992, Failures of the reduction principle in an ellsberg-type problem, *Theory and Decision* 32, 77-100.

Brenner, Menachem, and Yehuda Izhakian, 2011, Asset pricing and ambiguity: Empirical evidence, *Working Paper Stern School of Business.*

Bulan, Laarni T., 2005, Real options, irreversible investment and firm uncertainty: New evidence from U.S. firms, *Review of Financial Economics* 14, 255–279.

Camerer, Colin, and Martin Weber, 1992, Recent developments in modeling preferences: Uncertainty and ambiguity, *Journal of Risk and Uncertainty* 5, 325-370.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Dittmar, Robert F., 2002, Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns, *Journal of Finance* 57, 369-403.

Dow, James and Werlang, S. R. C., 1992, Excess volatility of stock prices and knightian uncertainty, *European Economic Review* 36, 631–638.

Drechsler, Itamar, 2013, Uncertainty, time-varying fear, and asset prices, *Journal of Finance* 68, 1843-1889.

Ehsani, Sina, Timothy Krause, and Donald Lien, 2013, An empirical index of knightian uncertainty, *Working Paper University of Texas at San Antonio*.

Ellsberg, Daniel, 1961, Risk, ambiguity, and the savage axioms, *Quarterly Journal of Economics* 75, 643-669.

Epstein, Larry G., and Martin Schneider, 2008, Ambiguity, infomation quality, and asset pricing, *Journal of Finance* 63, 197-228.

Epstein, Larry G., and Tan Wang, 1994, Intertemporal asset pricing under knightian uncertainty, *Econometrica* 62, 283–322.

Erbaş, S. Nuri, and Abbas Mirakhor, 2007, The equity premium puzzle, ambiguity aversion, and institutional quality, *IMF Working Paper*.

Fama, Eugene, 1965, The behavior of stock-market prices, *Journal of Business* 38, 34-105.

Fama, Eugene and James D. MacBeth, 1973, Risk, return, and equilibrium: empirical tests, *Journal of Political Economy* 81, 607-636.

Fama, Eugene and Kenneth French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fang, Hsing and Tsong-Yue Lai, 1997, Co-kurtosis and capital asset pricing, *Financial Review* 32, 293–307.

Garlappi, Lorenzo, Raman Uppal, and Tan Wang, 2007, Portfolio selection with parameter and model uncertain: A multi-prior approach, *Review of Financial Studies* 20, 41-81.

Gilboa, Itzhak and David Schmeidler, 1989, Maxmin expected utility with a non-unique prior, *Journal of Mathematical Economics* 18, 141–153.

Grullon, Gustave, Evgeny Lyandres, and Alexei Zhdanov, 2012, Real options, volatility, and stock returns, *Journal of Finance* 67, 1499-1537.

Guidolin, Massimo and Allan Timmermann, 2007, Asset allocation under multivariate regime switching, *Journal of Economic Dynamics and Control* 31, 3503-3544.

Harvey, Campbell R. and Akhtar Siddique, 1999, Autoregressive conditional skewness, *Journal of Financial and Quantitative Analysis* 34, 465–487.

Harvey, Campbell R. and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263–1295.

Izhakian, Yehuda, 2012, Shadow probability theory and ambiguity measurement, *Working Paper New York University.*

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jeong, Daehee, Hwagyun Kim, and Joon Y. Park, 2015, Does ambiguity matter? Estimating asset pricing models with a multiple-priors recursive utility, *Journal of Financial Economics* 115, 361-382.

Jiang, Guohua, Charles M. C. Lee, and Grace Y. Zhang, 2005, Information uncertainty and expected returns, *Review of Accounting Studies* 10, 185-221.

Ju, Nengjiu and Jianjun Miao, 2012, Ambiguity, learning, and asset returns, *Econometrica* 80, 559-591.

Kacperczyk, Marcin and Paul Damien, 2011, Asset allocation under distribution uncertainty, *Working Paper University of Texas at Austin.*

Kimball, Miles S., 1993, Standard risk aversion, *Econometrica* 61, 589-611.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2005, A smooth model of decision making under ambiguity, *Econometrica* 73, 1849-1892.

Knight, Frank H., 1921. Risk, uncertainty and profit, 1st edn, Houghton Mifflin, Boston, MA.

Kraus, Alan and Robert H. Litzenberger, 1976, Skewness preference and the valuation of risk assets, *Journal of Finance* 31, 1085-1100.

Lehmann, B., 1990, Fads, martingales, and market efficiency, *Quarterly Journal of Economics* 105, 1–28.

Levy, Haim, 1969, A Utility function depending on the first three moments, *Journal of Finance* 24, 715-719.

Liu, Jun, Francis A. Longsta, and Jun Pan, 2003, Dynamic asset allocation with event risk, *Journal of Finance* 58, 231-259.

Liu, Jun, Jun Pan, and Tan Wang, 2005, An equilibrium model of rare event premia and its implication for option smirks, *Review of Financial Studies* 18, 131-164.

Markowitz, Harry, 1952, Portfolio selection, *Journal of Finance* 7, 77-91.

Olsen, Robert A. and George H. Troughton, 2000, Are risk premium anomalies caused by ambiguity?, *Financial Analysts Journal* 556, 24-31.

Pastor, Lubos and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642-685.

Pratt, John W. and Richard J. Zeckhauser, 1987, Proper risk aversion, *Econometrica* 55, 143-154.

Rosenberg, Barr, 1974, Extra-market components of covariance in security returns, *Journal of Financial and Quantitative Analysis* 9, 263-274.

Sherman, Roger, 1974, The psychological difference between ambiguity and risk, *Quarterly Journal of Economics* 88, 166-169.

Smith, Daniel R., 2007, Conditional coskewness and asset pricing, *Journal of Empirical Finance* 14, 91–119.

Tsay, Ruey S., 2010. Analysis of financial time series, 3rd edn, John Wiley & Sons, Inc., New York, NY.

Welch, Ivo, 2000, Views of financial economists on the equity premium and on professional controversies, *Journal of Business* 73, 501-537.

Yates, J. Frank and Lisa G Zukowski, 1976, Characterization of ambiguity in decision making, *Behavioral Science* 21, 19-25.

Zhang, X. Frank, 2006, Information uncertainty and stock returns, *Journal of Finance* 61, 105-136.

Table I
Descriptive Statistics

This table reports summary statistics and correlation coefficients for the key variables. DD is the average of previous 24 EMDs (the Earth Mover's Distance). Each EMD non-parametrically measures the difference between two probability distributions using previous 12 months. SIZE is the natural logarithm of the market value of equity of the company (in thousands of dollars) measured by times series average of a firm's market capitalization for the most recent 12 months. BM is the book value of equity divided by its market value at the end of the last fiscal year. BETA is estimated by regressing each stock's daily return on the value weighted index return using the previous year's data. MOM is calculated as the return over the 11 months prior to that month. REV is the previous month’s stock return. ILLIQ by Amihud (2002) is measured as the ratio of the absolute monthly stock return to its dollar trading volume for each stock, scaled by 1,000. IDIOVOL is calculated as the standard deviation of the daily residuals in month t from the CAPM. SKEW is the historical third-order centralized moment using daily returns within year t. KURT is the fourth-order centralized moment. MAX is defined by the maximum daily return over the past month. The sample includes all firms listed in NYSE, AMEX, and NASDAQ from 1965 to 2012.

Panel A: Descriptive Statistics

	DD	SIZE	BM	BETA	MOM	REV	ILLIQ	IDIOVOL	SKEW	KURT	MAX
N	2717967	3401816	2324268	3095031	3118816	3404580	3098907	3118727	3122055	3122055	3404580
Mean	0.0251	11.3702	0.8364	0.8540	0.1341	0.0112	0.0026	6.3253	0.5646	10.3961	0.0742
P1	0.0159	7.0135	0.0225	-1.0283	-0.8177	-0.3846	0.0000	0.9979	-2.7734	2.9125	0.0000
P5	0.0207	8.0693	0.1204	-0.2068	-0.6023	-0.2178	0.0000	1.7335	-0.9265	3.3843	0.0104
P10	0.0216	8.6874	0.1970	0.0062	-0.4582	-0.1500	0.0000	2.2568	-0.4330	3.7212	0.0164
P25	0.0223	9.8287	0.3741	0.3314	-0.2000	-0.0635	0.0000	3.3500	0.0118	4.5246	0.0286
median	0.0230	11.2507	0.6581	0.7896	0.0517	0.0000	0.0001	5.1643	0.3865	6.1741	0.0504
P75	0.0242	12.7895	1.0770	0.7896	0.3153	0.0676	0.0007	7.9326	0.8955	10.1643	0.0889
P90	0.0295	14.2042	1.6654	1.8452	0.7005	0.1667	0.0030	11.6216	1.7306	19.2570	0.1500
P95	0.0362	15.0769	2.1874	2.2320	1.0804	0.2593	0.0066	14.6321	2.6401	30.1461	0.2093
P99	0.0617	16.7713	3.6054	3.1289	2.4706	0.5667	0.0275	22.7629	6.0196	75.8818	0.4068
Stdev	0.0114	2.1436	0.6978	0.8154	0.7536	0.1759	0.2807	4.7530	1.4131	14.5901	0.0962

Panel B: Correlation Matrix (Pearson Correlations Are Shown above the Diagonal with Spearman Below)

	DD	SIZE	BM	BETA	MOM	REV	ILLIQ	IDIOVOL	SKEW	KURT	MAX
DD	1.0000	-0.1806	0.0053	-0.0012	0.0645	0.0313	0.0045	0.4788	0.2455	0.2773	0.2890
SIZE	-0.2308	1.0000	-0.3457	0.1953	0.0798	0.0282	-0.0265	-0.4117	-0.2209	-0.1560	-0.2765
BM	-0.0348	-0.3576	1.0000	-0.1051	-0.1622	-0.0772	0.0153	0.0364	0.0132	0.0192	0.0531
BETA	0.0542	0.2305	-0.1507	1.0000	0.0548	-0.0011	-0.0114	0.1358	-0.0283	-0.0547	0.0456
MOM	-0.0894	0.1849	-0.2243	0.0014	1.0000	-0.0049	-0.0042	0.1200	0.1794	0.0422	-0.0448
REV	-0.0303	0.0848	-0.0834	-0.0100	0.0189	1.0000	0.0031	0.0890	0.1081	0.0327	0.3239
ILLIQ	0.0270	-0.6749	0.3011	-0.2311	-0.0701	-0.0041	1.0000	0.0044	0.0032	0.0171	0.0065
IDIOVOL	0.4460	-0.4870	-0.0226	0.2002	-0.1437	-0.0451	0.1656	1.0000	0.3319	0.2832	0.5484
SKEW	0.1782	-0.2963	0.0095	-0.0129	0.1959	0.0589	0.1984	0.3104	1.0000	0.4760	0.2284
KURT	0.2750	-0.2066	0.0157	-0.0483	-0.0741	-0.0255	0.1004	0.2735	0.3336	1.0000	0.1561
MAX	0.3108	-0.3494	-0.0116	0.1515	-0.1857	0.2015	0.1473	0.6642	0.2246	0.1001	1.0000

Table II
Descriptive statistics based on Difference of Distributions

This table reports summary statistics and correlation coefficients for the key variables. DD is the average of previous 24 EMDs (the Earth Mover's Distance). Each EMD non-parametrically measures the difference between two probability distributions using previous 12 months. The decile portfolios updated each month are formed by the sizes of DD statistics estimated using daily demeaned individual stock return over previous 36 months. Portfolio 'S' is the portfolio of stocks with the lowest DD, Portfolio 'B' is the portfolio of stocks with the highest DD. SIZE is the natural logarithm of the market value of equity of the company (in thousands of dollars) measured by times series average of a firm's market capitalization for the most recent 12 months. BM is the book value of equity divided by its market value at the end of the last fiscal year. BETA is estimated by regressing each stock's daily return on the value weighted index return using the previous year's data. MOM is calculated as the return over the 11 months prior to that month. REV is the previous month's stock return. ILLIQ by Amihud (2002) is measured as the ratio of the absolute monthly stock return to its dollar trading volume for each stock, scaled by 1,000. IDIOVOL is calculated as the standard deviation of the daily residuals in month t from the CAPM. SKEW is the historical third-order centralized moment using daily returns within year t. KURT is the fourth-order centralized moment. MAX is defined by the maximum daily return over the past month. AGE is the natural logarithm of the firm age, and REAL OPTION_D is a dummy variable for real options with the value of 1 if a firm belongs to high-tech, pharmaceutical, and biotechnology industries, or 0 otherwise. The sample includes all firms listed in NYSE, AMEX, and NASDAQ from 1965 to 2012.

DD	DD	SIZE	BM	BETA	MOM	REV	ILLIQ	IDIOVOL	SKEW	KURT	MAX	AGE	REAL OPTION_D
S	0.0199	1,428,190	0.8840	0.5809	0.1168	0.0089	0.0030	3.2834	0.3674	9.3751	0.0362	2.8436	0.0534
2	0.0219	1,715,206	0.8641	0.8882	0.1449	0.0108	0.0022	4.6253	0.4046	7.2400	0.0521	2.6611	0.1005
3	0.0223	1,472,392	0.8642	0.9270	0.1503	0.0111	0.0021	4.7840	0.4117	7.1026	0.0543	2.6473	0.1082
4	0.0226	1,481,521	0.8698	0.9266	0.1494	0.0112	0.0019	4.7964	0.4114	7.1709	0.0546	2.6681	0.1067
5	0.0229	1,486,758	0.8721	0.9241	0.1449	0.0115	0.0022	4.8098	0.4124	7.3879	0.0550	2.6817	0.1050
6	0.0232	1,403,613	0.8738	0.9230	0.1377	0.0114	0.0020	4.9351	0.4172	7.7766	0.0567	2.6898	0.1074
7	0.0236	1,283,590	0.8808	0.9330	0.1373	0.0117	0.0021	5.3640	0.4600	8.6329	0.0623	2.6712	0.1176
8	0.0245	876,229	0.9047	0.9428	0.1390	0.0126	0.0025	6.3666	0.5759	10.2045	0.0755	2.5900	0.1432
9	0.0268	397,942	0.9372	0.9473	0.1395	0.0145	0.0055	8.2368	0.8294	13.4310	0.0994	2.4638	0.1670
B	0.0407	130,811	0.9457	0.9121	0.2011	0.0243	0.0058	11.8670	1.5418	20.9803	0.1440	2.3134	0.1716

Table III
Portfolio Returns Sorted on DD

This table presents value-weighted and equal-weighted average monthly returns for portfolios formed on DD within a month. DD is the average of previous 24 EMDs (the Earth Mover's Distance). Each EMD non-parametrically measures the difference between two probability distributions using previous 12 months. The decile portfolios updated each month are formed by the sizes of DD statistics estimated using daily demeaned individual stock return over previous 36 months. Portfolio ‘S’ is the portfolio of stocks with the lowest DD, Portfolio ‘B’ is the portfolio of stocks with the highest DD, ‘S-B’ is the return difference between the low DD and high DD portfolios, and *t*-statistics are reported in parentheses. The sample includes all firms listed in NYSE, AMEX, and NASDAQ from 1965 to 2012.

	DD	Value-weighted	Equal-weighted
S	0.0199	0.7960	1.0988
2	0.0219	0.9768	1.4251
3	0.0223	1.0090	1.4600
4	0.0226	0.9987	1.4395
5	0.0229	1.0305	1.4868
6	0.0232	0.9496	1.5027
7	0.0236	0.9564	1.5867
8	0.0245	1.1292	1.9503
9	0.0268	1.4161	2.5644
B	0.0407	1.5166	4.6755
B-S		0.7205	3.5766
t(B-S)		(3.06)	(10.91)

Table IV
Portfolio Returns by Various Estimation and Holding Periods

This table reports the value-weighted average returns for portfolios formed on DD using various estimation periods (*J*) and holding periods (*K*). DD is the average of previous 24 EMDs (the Earth Mover's Distance). Each EMD non-parametrically measures the difference between two probability distributions using previous 12 months. The decile portfolios updated each month are formed by the sizes of DD statistics estimated using daily demeaned individual stock return over previous 24 months and 36 months (*J*). We report the time series average of portfolios over the next one, three, six, and twelve months (*K*). Portfolio 'S' is the portfolio of stocks with the lowest DD, Portfolio 'B' is the portfolio of stocks with the highest DD, 'S-B' is the return difference between the low DD and high DD portfolios, and *t*-statistics are reported in parentheses. The sample includes all firms listed in NYSE, AMEX, and NASDAQ from 1965 to 2012.

	J=24/*K*=3	*J*=24/*K*=6	*J*=24/*K*=9	*J*=24/*K*=12
S	2.6394	5.3913	8.1571	10.8121
2	2.9094	5.9287	9.0073	12.0383
3	2.9010	5.9677	9.1179	12.2606
4	3.0511	6.1821	9.2703	12.2800
5	3.1809	6.2735	9.2781	12.1531
6	2.9632	6.0821	9.1604	12.4142
7	2.7667	5.7306	8.8804	12.1410
8	3.4525	7.2408	10.7424	14.1088
9	4.1437	8.1928	12.5004	16.8899
B	4.6124	9.4979	14.8108	19.9695
B-S	1.9729	4.1066	6.6537	9.1574
T(B-S)	(4.52)	(6.71)	(8.38)	(9.34)
	J=36/K=3	*J=36/K=6*	*J=36/K=9*	*J=36/K=12*
S	2.5190	5.2174	7.9627	10.6022
2	2.8789	5.8725	8.7253	11.8721
3	2.9533	5.8304	9.1354	12.4148
4	3.0319	6.1993	9.3224	12.2556
5	3.2945	6.6560	9.9253	12.9723
6	2.9520	6.1281	9.4165	12.7803
7	2.9933	6.3290	9.4562	12.7738
8	3.4409	6.8257	10.4358	13.9810
9	4.3556	8.8714	13.2567	17.5006
B	4.9849	10.1116	15.5418	21.0326
B-S	2.4659	4.8942	7.5791	10.4304
t(B-S)	(5.54)	(7.49)	(9.25)	(10.16)

Table V
Portfolios Returns Sorted on DD and Firm Characteristics

This table reports value-weighted and equal-weighted average monthly returns for portfolios based on DD and firm characteristics. DD is the average of previous 24 EMDs (the Earth Mover's Distance). Each EMD non-parametrically measures the difference between two probability distributions using previous 12 months. In each case, we first sort the stocks into deciles using the firm characteristic. Within each characteristic decile, we sort stocks into ten additional portfolios based on DD and compute the returns on the corresponding portfolios over the subsequent month. This table presents average returns across the firm characteristic deciles. Portfolio 'S' is the portfolio of stocks with the lowest DD, Portfolio 'B' is the portfolio of stocks with the highest DD, 'S-B' is the return difference between the low DD and high DD portfolios, and *t*-statistics are reported in parentheses. BETA is estimated by regressing each stock's daily return on the value weighted index return using the previous year's data. SIZE is the natural logarithm of the market value of equity of the company (in thousands of dollars) measured by times series average of a firm's market capitalization for the most recent 12 months. BM is the book value of equity divided by its market value at the end of the last fiscal year. MOM is calculated as the return over the 11 months prior to that month. REV is the previous month's stock return. ILLIQ by Amihud (2002) is measured as the ratio of the absolute monthly stock return to its dollar trading volume for each stock, scaled by 1,000. Each characteristic is defined in the Appendix. The sample includes all firms listed in NYSE, AMEX, and NASDAQ from 1965 to 2012.

Panel A. Value-weighted

	BETA	SIZE	BM	MOM	REV	ILLIQ
S	0.9619	1.2468	1.1318	0.9922	0.9677	0.9402
2	0.9492	1.5529	1.1720	0.8706	1.0162	1.0928
3	0.9951	1.6888	1.2242	0.9859	1.0377	1.1024
4	1.0524	1.7109	1.3327	0.9085	1.1283	1.0925
5	0.9860	1.8016	1.2538	0.9416	1.1571	1.0878
6	1.0324	1.8120	1.2947	0.9919	1.0269	1.0683
7	1.0868	1.9242	1.3524	1.0074	1.2763	1.1833
8	1.1343	1.9540	1.4355	1.1520	1.2813	1.3253
9	1.3654	2.0907	1.6119	1.2466	1.4883	1.5314
B	1.6169	2.3133	1.9513	1.6617	1.6877	1.9448
B-S	0.6550	1.0665	0.8194	0.6695	0.7200	1.0046
t(B-S)	(3.63)	(4.85)	(4.46)	(3.82)	(3.74)	(4.76)

Panel B. Equal-weighted

	BETA	SIZE	BM	MOM	REV	ILLIQ
S	1.2682	1.2570	1.2422	1.2876	1.2622	1.1374
2	1.3649	1.5809	1.3970	1.3896	1.4316	1.4497
3	1.4390	1.7080	1.4175	1.5059	1.4968	1.4283
4	1.4547	1.7528	1.4558	1.5384	1.5573	1.4270
5	1.5132	1.8837	1.5126	1.5684	1.5953	1.5225
6	1.5577	1.9117	1.5099	1.6783	1.7028	1.5755
7	1.6666	2.0756	1.6433	1.7885	1.8291	1.7590
8	1.9387	2.1083	1.8081	2.0293	2.0555	2.0420
9	2.5383	2.3190	2.2018	2.4426	2.4146	2.6400
B	4.3783	2.6289	3.7817	4.0332	3.9213	4.1382
B-S	3.1102	1.3719	2.5395	2.7455	2.6591	3.0007
t(B-S)	(11.80)	(5.98)	(11.38)	(11.68)	(10.69)	(9.60)

Table VI
Alphas of Portfolios Sorted on DD

This table reports the alphas of the CAPM, the Fama-French 3-factor model, and the Carhart (1997) 4-factor models for 10 portfolios based on DD. DD is the average of previous 24 EMDs (the Earth Mover's Distance). Each EMD non-parametrically measures the difference between two probability distributions using previous 12 months. Alphas are from a time series regression of the monthly returns on Rm-Rf, SMB, HML, and UMD as in Fama and French (1993) and Carhart (1997). The decile portfolios updated each month are formed by the sizes of DD statistics estimated using daily demeaned individual stock return over previous 36 months. Portfolio 'S' is the portfolio of stocks with the lowest DD, Portfolio 'B' is the portfolio of stocks with the highest DD, 'S-B' is the return difference between the low DD and high DD portfolios, and *t*-statistics are reported in parentheses. The sample includes all firms listed in NYSE, AMEX, and NASDAQ from 1965 to 2012.

Panel A. Value-weighted

	CAPM		Fama-French 3 Factor		Carhart 4 Factor	
	Alpha	Adj Rsq	Alpha	Adj Rsq	Alpha	Adj Rsq
S	0.0340	0.7842	0.0001	0.8007	0.0179	0.8008
2	0.1251	0.9276	0.1072	0.9278	0.1128	0.9277
3	0.1326	0.9124	0.1223	0.9125	0.1216	0.9123
4	0.1320	0.9293	0.0998	0.9307	0.0823	0.9308
5	0.1628	0.9143	0.1205	0.9208	0.1513	0.9216
6	0.0771	0.9073	0.0636	0.9112	0.1004	0.9123
7	0.0699	0.9033	0.0703	0.9029	0.1357	0.9065
8	0.2069	0.8384	0.1989	0.8517	0.2298	0.8521
9	0.4358	0.7243	0.3899	0.7803	0.4516	0.7817
B	0.5229	0.6214	0.3879	0.7808	0.4759	0.7835
B-S	0.4889		0.3878		0.4580	
t(B-S)	(2.27)		(2.40)		(2.85)	

Panel B. Equal-weighted

	CAPM		Fama-French 3 Factor		Carhart 4 Factor	
	Alpha	Adj Rsq	Alpha	Adj Rsq	Alpha	Adj Rsq
S	0.3674	0.6630	0.1647	0.7757	0.2342	0.7824
2	0.5586	0.7789	0.2866	0.9217	0.3615	0.9262
3	0.5765	0.7706	0.2861	0.9243	0.3591	0.9282
4	0.5553	0.7807	0.2705	0.9295	0.3453	0.9337
5	0.5952	0.7886	0.3073	0.9340	0.3932	0.9394
6	0.6139	0.7877	0.3372	0.9351	0.4584	0.9461
7	0.6756	0.7674	0.4104	0.9254	0.5867	0.9461
8	1.0133	0.6909	0.7479	0.8782	0.9608	0.9026
9	1.5882	0.5628	1.2600	0.7779	1.5466	0.8089
B	3.6738	0.3635	3.3154	0.5696	3.6477	0.5938
B-S	3.3064		3.1507		3.4136	
t(B-S)	(10.80)		(11.66)		(12.86)	

Table VII
Firm-Level Cross-Sectional Return Regressions with DD and firm characteristics

Each month, we run a firm-level cross-sectional regression of the one-month ahead return on DD and six control variables in the month. DD is the average of previous 24 EMDs (the Earth Mover's Distance). Each EMD non-parametrically measures the difference between two probability distributions using previous 12 months. BETA is estimated by regressing each stock's daily return on the value weighted index return using the previous year's data. SIZE is the natural logarithm of the market value of equity of the company (in thousands of dollars) measured by times series average of a firm's market capitalization for the most recent 12 months. BM is the book value of equity divided by its market value at the end of the last fiscal year. MOM is calculated as the return over the 11 months prior to that month. REV is the previous month's stock return. ILLIQ by Amihud (2002) is measured as the ratio of the absolute monthly stock return to its dollar trading volume for each stock, scaled by 1,000. Each characteristic is defined in the Appendix. This table presents the time series average of the cross-sectional regression coefficients and corresponding *t*-statistics following methodology of Fama and MacBeth (1973). The sample includes all firms listed in NYSE, AMEX, and NASDAQ from 1965 to 2012.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DD	0.3851	0.3507	0.3424	0.1889	0.3542	0.3991	0.3904	0.2557
	(3.09)	(2.97)	(2.84)	(2.18)	(3.04)	(3.26)	(3.19)	(3.16)
BETA		0.0002						0.0015
		(0.17)						(1.35)
BM			0.0128					0.0118
			(17.18)					(21.31)
SIZE				-0.0018				-0.0008
				(-4.62)				(-2.11)
MOM					0.0044			0.0073
					(2.59)			(5.76)
REV						-0.0555		-0.0589
						(-12.89)		(-16.65)
ILLIQ							0.1143	-0.15376
							(1.39)	(-2.37)
AdjRSQ	0.0062	0.0241	0.0157	0.0209	0.0184	0.0168	0.0080	0.0603

Table VIII
Portfolios Returns Sorted on DD and Return Distribution Characteristics

This table reports value-weighted and equal-weighted average monthly returns for portfolios based on DD and the return distribution characteristics. DD is the average of previous 24 EMDs (the Earth Mover's Distance). Each EMD non-parametrically measures the difference between two probability distributions using previous 12 months. In each case, we first sort the stocks into deciles using the return distribution characteristic. Within each characteristic decile, we sort stocks into ten additional portfolios based on DD and compute the returns on the corresponding portfolios over the subsequent month. This table presents average returns across the firm characteristic deciles. Portfolio 'S' is the portfolio of stocks with the lowest DD, Portfolio 'B' is the portfolio of stocks with the highest DD, 'S-B' is the return difference between the low DD and high DD portfolios, and *t*-statistics are reported in parentheses. IDIOVOL is calculated as the standard deviation of the daily residuals in month t from the CAPM. SKEW is the historical third-order centralized moment using daily returns within year t. KURT is the fourth-order centralized moment. MAX is defined by the maximum daily return over the past month. Each characteristic is defined in the Appendix. The sample includes all firms listed in NYSE, AMEX, and NASDAQ from 1965 to 2012.

Panel A. Value-weighted				
	IDIOVOL	SKEWS	KURT	MAX
S	1.0757	0.9038	0.9428	0.9338
2	1.0266	1.0603	0.9953	0.9850
3	1.1517	1.0402	1.0470	1.0022
4	1.1103	1.0863	1.0659	1.0600
5	1.0748	1.0998	1.0393	1.1229
6	1.0925	1.1336	1.1088	1.0766
7	1.2222	1.0369	1.0383	1.1082
8	1.2115	1.1222	1.1742	1.2744
9	1.2822	1.2251	1.1972	1.2403
B	1.5098	1.6049	1.5531	1.6548
B-S	0.4341	0.7012	0.6103	0.7210
t(B-S)	(4.66)	(3.32)	(2.75)	(5.40)
Panel B. Equal-weighted				
	IDIOVOL	SKEWS	KURT	MAX
S	1.5688	1.1598	1.1137	1.3482
2	1.6515	1.3890	1.3638	1.4355
3	1.7391	1.4393	1.4030	1.4943
4	1.7212	1.4579	1.4447	1.6349
5	1.7981	1.5007	1.4944	1.7138
6	1.9022	1.5508	1.4932	1.7872
7	1.9650	1.6929	1.6350	1.9243
8	2.0386	1.9308	1.9064	2.1395
9	2.2184	2.6369	2.5170	2.4614
B	2.6759	4.3965	4.7590	3.3339
B-S	1.1071	3.2367	3.6453	1.9857
t(B-S)	(12.23)	(11.48)	(11.78)	(11.71)

Table IX

Firm-Level Cross-Sectional Return Regressions with DD and Return Distribution characteristics

Each month, we run a firm-level cross-sectional regression of the one-month ahead return on DD and return distribution characteristics as well as six control variables in the month. DD is the average of previous 24 EMDs (the Earth Mover's Distance). Each EMD non-parametrically measures the difference between two probability distributions using previous 12 months. IDIOVOL is calculated as the standard deviation of the daily residuals in month t from the CAPM. SKEW is the historical third-order centralized moment using daily returns within year t. SSKEW is the systematic skewness of stock i in year t and ISKEW is the idiosyncratic skewness of stock i in year t. KURT is the fourth-order centralized moment. MAX is defined by the maximum daily return over the past month. Each characteristic is defined in the Appendix. This table presents the time series average of the cross-sectional regression coefficients and corresponding *t*-statistics following methodology of Fama and MacBeth (1973). The sample includes all firms listed in NYSE, AMEX, and NASDAQ from 1965 to 2012.

	(1)	(2)	(3)	(4)	(5)	(6)
DD	0.2630	0.2891	0.2684	0.2626	0.3197	0.2719
	(3.86)	(3.57)	(3.35)	(3.86)	(3.78)	(3.60)
BETA	0.0012	0.0015	0.0015	0.0012	0.0014	0.0019
	(1.27)	(1.38)	(1.30)	(1.29)	(1.30)	(1.84)
BM	0.0117	0.0118	0.0117	0.0117	0.0118	0.0117
	(22.14)	(21.43)	(21.53)	(22.15)	(21.34)	(21.69)
SIZE	-0.0007	-0.0009	-0.0008	-0.0007	-0.0008	-0.0008
	(-2.36)	(-2.29)	(-2.18)	(-2.36)	(-2.16)	(-2.40)
MOM	0.0075	0.0076	0.0075	0.0074	0.0074	0.0073
	(5.94)	(5.90)	(5.99)	(5.94)	(5.81)	(5.89)
REV	-0.0616	-0.0585	-0.0602	-0.0616	-0.0587	-0.0580
	(-17.48)	(-16.56)	(-17.08)	(-17.47)	(-16.64)	(-15.33)
ILLIQ	-0.1386	-0.1551	-0.1554	-0.1383	-0.1447	-0.1601
	(-2.37)	(-2.39)	(-2.43)	(-2.37)	(-2.24)	(-2.51)
IDIOVOL	0.0020					
	(0.09)					
SKEW		-0.00054				
		(-2.79)				
SSKEW			-0.0048			
			(-1.46)			
ISKEW				0.0024		
				(0.11)		
KURT					-0.0001	
					(-5.01)	
MAX						-0.01487
						(-1.97)
AdjRSQ	0.0671	0.0609	0.0627	0.0671	0.0610	0.0639